REVUE BELGE DE STATISTIQUE ET DE RECHERCHE OPERATIONNELLE

Vol. 9 - N° 4
SEPTEMBRE 1969

BELGISCH TIJDSCHRIFT VOOR STATISTIEK EN OPERATIONEEL ONDERZOEK

Vol. 9 - Nr 4
SEPTEMBER 1969

La « Revue Belge de Statistique et de Recherche Opérationnelle » est publiée avec l'appui du Ministère de l'Education nationale et de la Culture, par les Sociétés suivantes :

SOGESCI. — Société Belge pour l'Application des Méthodes scientifiques de Gestion.
Secrétariat : 66, rue de Neufchâtel, Bruxelles 6. Tél. 37.19.76.

S.B.S. — Société Belge de Statistique.
Siège social : 44, rue de Louvain, Bruxelles.
Secrétariat : 44, rue de Louvain, Bruxelles.

Het « Belgisch Tijdschrift voor Statistiek en Operationeel Onderzoek » wordt uitgegeven met de steun van het Ministerie van Nationale Opvoeding en Cultuur, door de volgende Verenigingen :

SOGESCI. — Belgische Vereniging voor Toepassing van Wetenschappelijke Methodes in het Bedrijfsbeheer.
Secretariaat : Neufchâtelstraat 66, Brussel 6. Tel. 37.19.76.

S.B.S. — Belgische Vereniging voor Statistiek.
Maatschappelijke zetel : 44, Leuvensestraat, Brussel.
Secretariaat : 44, Leuvensestraat, Brussel.

Comité de Direction

J. FEDERWISCH, Ingénieur civil, 127, avenue L. Wiener, Bruxelles 17.
S. MORNARD, Licencié en Sciences, 51, rue Souveraine, Bruxelles 5.
R. SNEYERS, Docteur en Sciences, Chef de Travaux à l'Institut Royal Météorologique de Belgique, 3, avenue Circulaire, Bruxelles 18.

Comité de Screening

F. LAMBERT, Professeur à la Faculté Polytechnique de Mons, 18, avenue Général Longueville, Bruxelles 15.
R. SNEYERS, Docteur en Sciences, Chef de Travaux à l'Institut Royal Météorologique de Belgique, 3, avenue Circulaire, Bruxelles 18.

Rédaction

R. SNEYERS, Docteur en Sciences, Chef de Travaux à l'Institut Royal Météorologique de Belgique, 3, avenue Circulaire, Bruxelles 18.

Secrétariat

J.H. LENTZEN, 66, rue de Neufchâtel, Bruxelles 6 - Tél. 37.19.76.

Directie Comité

J. FEDERWISCH, Burgerlijk Ingenieur, 127, L. Wienerlaan, Brussel 17.
S. MORNARD, Lic. in de Wetenschappen, 51, Souverainestraat, Brussel 5.
R. SNEYERS, Dr in de Wetenschappen, Werkleider bij het Koninklijk Meteorologisch Instituut van België, 3, Ringlaan, Brussel 18.

Screening Comité

F. LAMBERT, Professor bij de Faculté Polytechnique de Mons, Gl de Longuevillelaan, Brussel 15.
R. SNEYERS, Dr in de Wetenschappen, Werkleider bij het Koninklijk Meteorologisch Instituut van België, 3, Ringlaan, Brussel 18.

Redactie

R. SNEYERS, Dr in de Wetenschappen, Werkleider bij het Koninklijk Meteorologisch Instituut van België, 3, Ringlaan, Brussel 18.

Secretariaat

J.H. LENTZEN, 66 Neufchâtelstraat, Brussel 6 - Tel. 37.19.76.

REVUE BELGE DE STATISTIQUE
ET DE RECHERCHE OPERATIONNELLE

VOL. 9 - N° 4 - SEPTEMBRE 1969 VOL. 9 - Nr 4 - SEPTEMBER 1969

SOMMAIRE — INHOUD

BELGISCH TIJDSCHRIFT VOOR STATISTIEK
EN OPERATIONEEL ONDERZOEK

COMMENTS ON THE ASSUMPTION OF NORMALITY INVOLVED IN THE USE OF SOME SIMPLE STATISTICAL TECHNIQUES

by E.S. PEARSON
University College London

1. Introduction.

Statistical text books and manuals providing guidance for the user of statistical techniques in industrial quality control are accustomed to warn their readers that certain methods of analysis and inference should only be used if the distributions of the observed variables are normal or approximately normal. While writers on theoretical statistics have for many years devoted a good deal of attention to the problem of the so-called robustness of the more important statistical techniques, the results of these investigations do no so far seem to have been presented in a way which helps the user to form his own judgement on the risk he is taking in assuming normality in variation. The difficulty in simple presentation of the position is certainly considerable, for three elements enter into the problem:

(a) the form of departure from normality in the parent population sampled;

(b) the size of sample;

(c) the amount of error in a supposed significance level or confidence limit that the user is prepared to tolerate.

The following paper forms part of a memorandum which the author prepared recently for an International Standards Organisation Working Group on « The application of statistical techniques - Presentation of data ». It suggests one line of approach to this difficult problem.

2. Relation between robustness of tests and the form of distribution of the sampled population.

In what follows it is proposed to discuss what has been termed the « robustness » or « sensitivity to departure from normality » of the statistics u_α, t_α *and* χ_α^2 commonly used in drawing inferences in terms of probability on the basis of a single sample of n observations. We have particularly in mind the accuracy involved in the determination of, say, 100 $(1 - \alpha)\%$

upper confidence limits for an unknown mean and standard deviation through statements such as * :

$$\bar{X} \leqslant \bar{x} + u_{1-\alpha}\, \sigma/\sqrt{n} \qquad (1)$$

$$\bar{X} \leqslant \bar{x} + t_{1-\alpha}\, s/\sqrt{n} \qquad (2)$$

$$\sigma \leqslant \sqrt{\{\Sigma(x - \bar{x})^2\}}/\chi_\alpha \qquad (3)$$

where u is an N (0, 1) variable, t follows Student's distribution with $\nu = n - 1$ degree of freedom and χ^2 follows the chi-squared distribution, again with $\nu = n - 1$.

Experience has shown that the *shapes* of univariate frequency distributions may be very effectively classified in terms of the two moment ratios

$$\sqrt{\beta_1} = \gamma_1 = \mu_3/\sigma^3, \qquad \gamma_2 = \beta_2 - 3 = \mu_4/\sigma^4 - 3 ,$$

In Table 1 are given the values of these moment ratios calculated for a variety of data, drawn from industrial sources, which the writer has collected over a number of years. The fifteen points with co-ordinates (β_1, β_2) are plotted as open circles in figure 1. The square of $\gamma_1 = \sqrt{\beta_1}$, i.e. $\gamma_1^2 = \beta_1$ is used as co-ordinate because on this scaling certain important division lines in the field are represented exactly or approximately by straight lines. Although in the majority of cases several hundreds of observations, N, were available, it will be realised that the values of β_1 and β_2 are subject to varying amounts of sampling error.

While a number of the points cluster round the normal or Gaussian point for which $\beta_1 = 0$, $\beta_2 = 3$, we need still to ask in other cases what would be the error involved, for various sample sizes, if confidence limits derived from equations (1), (2) and (3) above were applied to samples from populations having the more divergent (β_1, β_2) points.

Approached in this way, the solution of the problem could be handled in two stages:

(a) The association of the (β_1, β_2) point of the population sampled with the percentage errors involved in using the « normal theory » factors in equations (1), (2) and (3).

(b) An identification by the user of the values of β_1, β_2 for the data from which he is sampling.

* Here u_P, t_ν, χ_P are values of each variable corresponding to a cumulative probability P.

Before considering (a) which is a problem in statistical sampling theory, a few comments on (b) are needed. If only a few measurements are available, perhaps only 10, 50 or even 100, it will be impossible to obtain a reliable estimate of the β_1, β_2 values for the sampled population, although asymmetry can often be detected visually. Here the graphical method may be very helpful in determining whether a transformation — e.g. by taking $\log x$, or $\sqrt{x}$ — will bring the data closer to symmetry.

Sometimes, however, it will have been possible to accumulate from past experience a considerable number of similar data. In this case, the calculation of the moment ratios of the frequency distribution presents little difficulty if appropriate mechanical equipment is available. For example, if the necessary programme card has been prepared, and the frequencies of a grouped distribution are fed successively by hand into an Olivetti Programma 101, the values of $\sqrt{\beta_1}$ and β_2 will be rapidly turned out and printed, the whole operation taking less than 5 minutes.

Consider now the theoretical problem involved in handling the stage (a) of the preceding paragraph. If we can determine the first four moments of the distributions of u, t and s/σ for samples of size n from any specified non-normal population, experience suggests that we could make a very good approximation to the percentage points of their sampling distributions. For example, we could assume that the distributions of these statistics are approximately represented by one of Karl Pearson's system of frequency curves having the correct first four moments, and hence obtain estimates of certain percentage points of their distributions by using the tables of Johnson *et al* (1963) which were prepared for use in just this kind of approximative procedure. As a result, we could get a rough idea of the errors involved in using the « normal theory » factors in determining confidence limits, given the population β_1 and β_2.

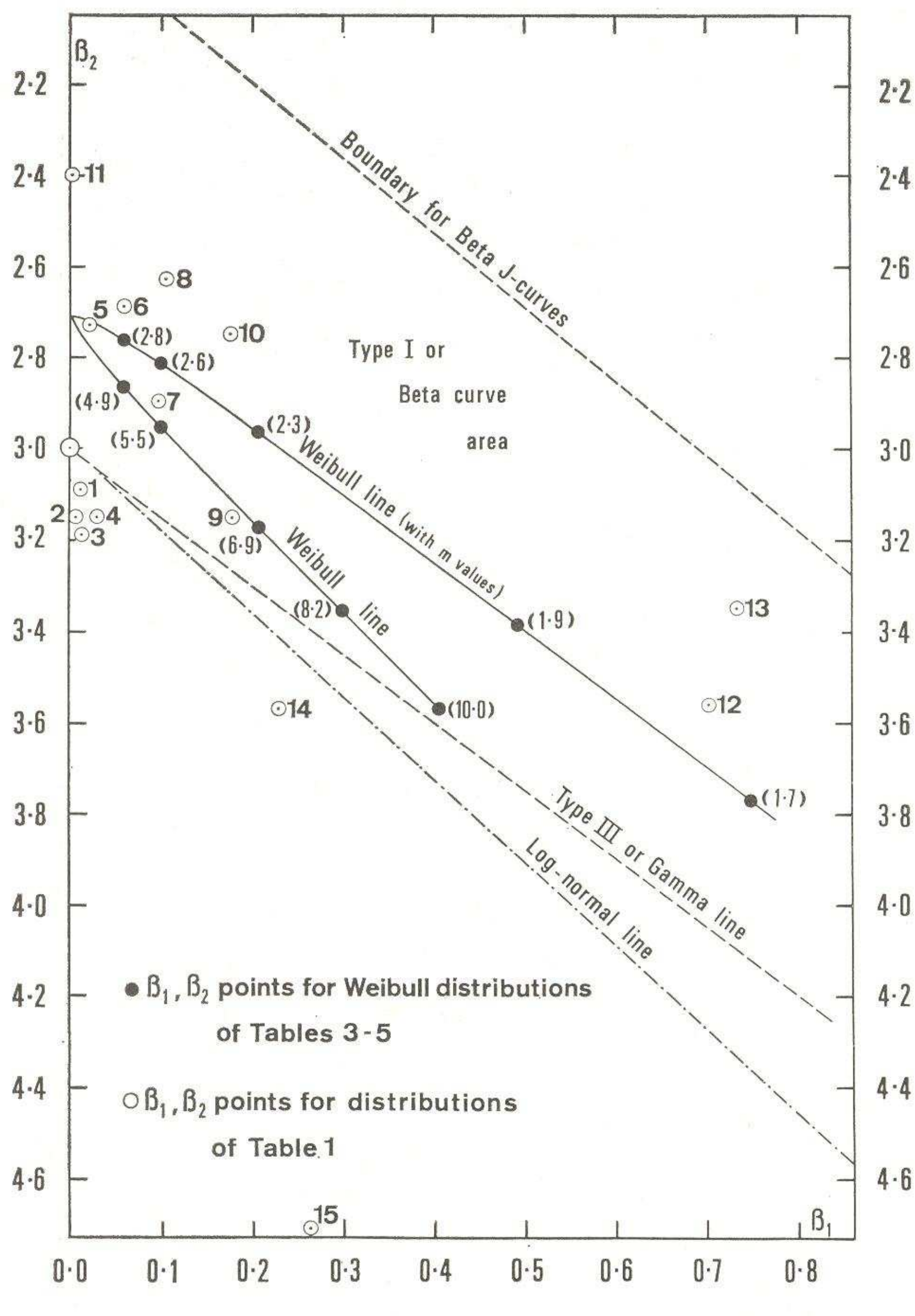
For u, the problem is straightforward; in fact if B_1, B_2 are the betas for a mean and β_1, β_2 the population values, then

$$B_1 = \beta_1/n, \qquad B_2 - 3 = (\beta_2 - 3)/n.$$

For $(s/\sigma)^2$ the B's can be calculated exactly, (see e.g. Kendall & Stuart, 1958, § 12.16) but depend on the 5^{th}, 6^{th} and 8^{th} moments or cumulants in the sampled population as well as on the 2^{nd}, 3^{rd} and 4^{th}. In the case of t the problem is more difficult; the values of $B_1(t)$, $B_2(t)$ have been expressed in terms of series expansions of the population moments or cumulants (see e.g. Geary, 1947, p. 214), but these series will generally not be convergent,

TABLE I. : Data regarding frequency distributions.

No.	N	$\sqrt{\beta_1}$	β_1	β_2	Material : characteristic measured	Source of data
1	849	0.101	0.010	3.09	Granular carbon : resistance in ohms.	Bell Telephone Laboratories
2	710	— 0.020	0.004	3.15	Minimum sound intensity : dynes per sq. in.	Bell Telephone Laboratories
3	995	0.115	0.013	3.19	Length of telephone conversations : time in seconds.	Bell Telephone Laboratories
4	1400	0.172	0.030	3.15	Telephone jacks : tension in ounces.	Bell Telephone Laboratories
5	180	— 0.141	0.020	2.73	Vacuum tubes : gain in d. b.	Bell Telephone Laboratories
6	1370	0.241	0.058	2.69	Telephone poles : depth of sapwood in ins.	Bell Telephone Laboratories
7	434	0.311	0.097	2.90	Malleable iron : yield strength in 100 lb. per sq. in.	Handbook of Metals, Vol. 1, 8th Ed. p. 370
8	300	0.324	0.105	2.63	Steel torson bar springs : life in 1000 cycles.	p. 219
9	71	0.424	0.179	3.15	Steel torson bar springs : life in 1000 cycles.	p. 219
10	1000	— 0.421	0.177	2.75	Duck cloth : warp strength in lb.	A.W. Bayes (1937). J. Roy Statist. Soc. Suppl., IV, p. 71
11	408	— 0.012	0.000	2.40	Malleable iron : elongation in 2 ins. per cent.	Handbook of Metals, Vol. 1 8th Ed. p. 370
12	238	0.838	0.702	3.56	Rotating bend fatigue tests : endurance to breaking point in 10^5 cycles.	Rolls Royce Aero Division
13	1000	0.856	0.733	3.35	Fibre strength in Indian cotton : strength in grms.	R.S. Koshal & A.J. Turner (1930). J. Text. Inst. XXXI, p. 325
14	790	0.478	0.229	3.57	Ferritic malleable iron : yield strength in 1000 lb. per sq. in.	Handbook of Metals, Vol. 1, 8th Ed., p. 370
15	2105	0.514	0.264	4.71	Sitka spruce : specific gravity.	U.S. Dept. of Agriculture, Tech. Bull. No 158.



β2
β1
2·2
2·4
2·6
2·8
3·0
3·2
3·4
3·6
3·8
4·0
4·2
4·4
4·6
0·0
0·1
0·2
0·3
0·4
0·5
0·6
0·7
0·8
Boundary for Beta J-curves
Type I or
Beta curve
area
Weibull line (with m values)
Weibull line
Type III or Gamma line
Log-normal line
(2·8)
(2·6)
(2·3)
(1·9)
(1·7)
(4·9)
(5·5)
(6·9)
(8·2)
(10·0)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
● β1, β2 points for Weibull distributions of Tables 3-5
○ β1, β2 points for distributions of Table 1

even for moderate values of the sample size, n. For this reason it seems likely that the only solution lies in Monte Carlo methods.

An extensive study of this kind into the behaviour of the t-distribution in samples from non-normal populations is being carried out by Fridshal & Posten at the University of Connecticut (Report No. 23 of Department of Statistics, 1966). In the meantime some very relevant and instructive results can be derived from a publication by Harter & Dubey (1967), which contains tables of percentage points for u, t and $v = s^2/\sigma^2$ for samples from Weibull distributions. These results are discussed in the following section. For t and v the Harter & Dubey tables were based on smoothed, Monte Carlo results.

3. Distribution of u, t and $(s/\sigma)^2$ for Weibull distributions.

The Weibull distribution of a random variable, x, is defined by the probability density distribution

$$f(x) = \frac{m}{b}\left(\frac{x-a}{b}\right)^{m-1} \exp\left\{-\left(\frac{x-a}{b}\right)^{m}\right\}, \qquad x \geqslant a. \tag{4}$$

A useful property of this distribution is that its probability integral or distribution function assumes the form

$$F(X) = \int_a^X f(x)\,dx = 1 - \exp\left\{-\left(\frac{x-a}{b}\right)^{m}\right\}, \tag{5}$$

the value of which may be readily obtained from tables of the negative exponential function.

The Shape of the Weibull distribution depends only on the parameter m. Values of the moments and moment ratios of the distribution for $m = 1.1\ (.1)\ 10.0$ are given in the Harter & Dubey Report. For $m > 3.6$, the distribution has negative skewness with longer tail to the left. But if we plot the values of $\beta_1 = \gamma_1^2$ and β_2 we get points on the two branches of the curve shown in figure 1. The region in the diagram in which these curves lie is that associated with the Type I or beta distributions of Karl Pearson's system of frequency curves,

$$f(x) = \{B(p, q)\}^{-1}\, x^{p-1}\,(1 - x)^{q-1}, \qquad 0 \leqslant x \leqslant 1. \tag{6}$$

That there is likely to be very little difference from the practical point of view between the shapes of the curves of types (4) and (6) which have the same values of the parameters β_1 and β_2 is suggested by the comparison made in Table 2. This shows for each system and for a selection of values of the parameters, m, of equation (4), the standardised distances from the mean to the lower and upper 5, and 1 per cent points. It is suggested that any conclusions drawn from Harter's Weibull results regarding the distributions of u, t and $(s/\sigma)^2$ will, broadly speaking, be applicable to samples from other unimodal distributions having similar values of β_1 and β_2.

Tables 3, 4 and 5 show for 10 different Weibull distributions the percentage errors involved in using the standard « normal theory » factors for u, t and χ^2/ν instead of the correct values as tabled by Harter & Dubey. The figures in the columns headed « % error » are the ratios:

$$100\ \{\text{per cent point tabled by Harter} - \text{normal theory value}\}$$

$$\div \text{ normal theory value.}$$

For example, in a sample of $n = 10$ the upper 5 per cent value for t on « normal theory » is 1.833, while for a Weibull distribution with $m = 1.9$, ($\beta_1 = 0.492$, $\beta_2 = 3.384$) it is 1.62. The « % error » is therefore taken as

$$100\ \{1.62 - 1.833\}/1.833 = -12\,.$$

For $m = 4.9$, 5.5, 6.9, 8.2 and 10.0, where the Weibull distributions are negatively skew, the upper and lower per cent points have been reversed so that errors in the factors for distributions on the two branches of the Weibull line shown in figure 1 may be compared. The following are some of the main points which emerge from a study of the tables.

The case of u; Tables 3 a and b.

(i) Errors associated with the corresponding lower and upper percent points are remarkably similar. As both have the same positive sign (for positive skewness in the parent population) the result is that the *width* of the confidence interval derived from the pair of limits

$$\bar{x} \pm 2.326\ \sigma/\sqrt{n} \qquad \text{or} \qquad \bar{x} \pm 1.645\ \sigma/\sqrt{n}$$

will be almost the same as for the limits given by Harter & Dubey, but there is a shift in the positive direction.

(ii) For a given population β_1 value, the errors on the two branches of the Weibull line are nearly the same.

TABLE 2. : Comparison of 5 % and 1 % standardised points of Weilbull and Beta distributions.

m		1.1		1.3		2.0		3.6		10.0	
$\sqrt{\beta_1}$		1.734		1.346		0.631		0.00056		— 0.638	
β_1		3.007		1.812		0.398		0.0000		0.407	
β_2		7.360		5.432		3.245		2.717		3.570	
		Weibull	Beta	Weibull	Beta	Weibull	Beta	Weibull	Beta	Weibull	Beta
Standardised deviates to	Lower 5 %	— 1.08	— 1.09	— 1.25^{-}	— 1.28	— 1.70	— 1.73	— 2.24	— 2.25^{-}	— 2.80	— 2.78
	Upper 1 %	— 1.02	— 1.03	— 1.15^{-}	— 1.15^{+}	— 1.42	— 1.42	— 1.66	— 1.65^{-}	— 1.82	— 1.81
	Upper 5 %	1.99	1.99	1.96	1.96	1.82	1.83	1.64	1.65^{-}	1.44	1.44
	Lower 1 %	3.47	3.47	3.23	3.24	2.72	2.73	2.26	2,25^{-}	1.87	1.84

TABLE 3a. : Errors in u_a.

m	β_1	β_2	n	Locer 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
				—	—	+	—	—	+	+	+	+	+	+	+
2.8	0.056	2.762	5	2.231	2.326	4.1	1.616	1.645	1.8	1.676	1.645	1.9	2.390	2.326	2.8
			10	2.263		2.7	1.624		1.3	1.667		1.3	2.374		2.1
			15	2.276		2.1	1.628		1.0	1.662		1.0	2.366		1.7
			20	2.283		1.8	1.630		0.9	1.660		0.9	2.362		1.5^+
			30	2.292		1.5^-	1.633		0.7	1.657		0.7	2.356		1.3
			60	2.303		1.0	1.636		0.5^+	1.654		0.5^+	2.348		0.9
				—	—	+	—	—	+	+	+	+	+	+	+
2.6	0.100	2.818	5	2.204	2.326	5.2	1.605	1.645	2.4	1.685	1.645	2.4	2.415	2.326	3.8
			10	2.244		3.5^+	1.617		1.7	1.673		1.7	2.392		2.8
			15	2.261		2.8	1.622		1.4	1.668		1.4	2.381		2.4
			20	2.270		2.4	1.625		1.2	1.665		1.2	2.374		2.1
			30	2.281		1.9	1.629		1.0	1.661		1.0	2.366		1.7
			60	2.295		1.3	1.633		0.7	1.656		0.7	2.355		1.2
				—	—	+	—	—	+	+	+	+	+	+	+
2.3	0.207	2.966	5	2.156	2.326	7.3	1.586	1.645	3.6	1.703	1.645	3.5^+	2.461	2.326	5.8
			10	2.211		4.9	1.603		2.6	1.686		2.5^-	2.424		4.4
			15	2.234		4.0	1.611		2.1	1.678		2.0	2.407		3.5^-
			20	2.247		3.4	1.616		1.8	1.674		1.8	2.397		3.1
			30	2.262		2.8	1.621		1.5^-	1.668		1.4	2.385		2.5^+
			60	2.282		1.9	1.628		1.0	1.662		1.0	2.368		1.8
				—	—	+	—	—	+	+	+	+	+	+	+
1.9	0.492	3.384	5	2.072	2.326	10.9	1.551	1.645	5.7	1.732	1.645	5.3	2.541	2.326	9.2
			10	2.152		7.5^-	1.580		4.0	1.707		3.8	2.481		6.7
			15	2.186		6.0	1.592		3.2	1.695		3.0	2.454		5.5^+
			20	2.206		5.2	1.599		2.8	1.689		2.7	2.437		4.8
			30	2.229		4.2	1.608		2.2	1.681		2.2	2.418		4.0
			60	2.258		2.9	1.619		1.6	1.670		1.5^+	2.391		2.8
				—	—	+	—	—	+	+	+	+	+	+	+
1.7	0.748	3.772	5	2.017	2.326	13.3	1.527	1.645	1.2	1.750	1.645	6.4	2.594	2.326	11.5
			10	2.113		9.2	1.563		5.0^-	1.720		4.6	2.519		8.3
			15	2.115		7.4	1.579		4.0	1.707		3.8	2.485		6.8
			20	2.179		6.3	1.588		3.5^-	1.699		3.3	2.464		5.9
			30	2.207		5.1	1.599		2.8	1.689		2.7	2.439		4.9
			60	2.242		3.6	1.613		1.9	1.676		1.9	2.407		3.5^-

TABLE 3b. : Errors in u_{α}.

m	β_1	β_2	n	Lower 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
				—	—	+	—	—	+	+	+	+	+	+	+
4.9	0.058	2.865	5	2.236	2.326	3.9	1.613	1.645	1.9	1.677	1.645	1.9	2.395	2.326	3.0
			10	2.265		2.6	1.623		1.3	1.667		1.3	2.377		2.2
			15	2.277		2.1	1.627		1.1	1.663		1.1	2.369		1.8
			20	2.284		1.8	1.629		1.0	1.660		0.9	2.363		1.6
			30	2.292		1.5^-	1.632		0.8	1.657		0.7	2.357		1.3
			60	2.303		1.0	1.636		0.5^+	1.654		0.5^+	2.348		0.9
				—	—	+	—	—	+	+	+	+	+	+	+
5.5	0.101	2.957	5	2.211	2.326	4.9	1.603	1.645	2.6	1.687	1.645	2.6	2.422	2.326	4.1
			10	2.247		3.4	1.616		1.8	1.674		1.8	2.396		3.0
			15	2.263		2.7	1.621		1.5^-	1.668		1.4	2.384		2.5^-
			20	2.272		2.3	1.624		1.3	1.665		1.2	2.376		2.1
			30	2.282		1.9	1.628		1.0	1.661		1.0	2.368		1.8
			60	2.295		1.3	1.633		0.7	1.657		0.7	2.356		1.3
				—	—	+	—	—	+	+	+	+	+	+	+
6.9	0.207	3.172	5	2.169	2.326	6.7	1.584	1.645	3.7	1.703	1.645	3.5^+	2.470	2.326	6.2
			10	2.217		4.7	1.603		2.6	1.686		2.5^-	2.429		4.4
			15	2.237		3.8	1.611		2.1	1.678		2.0	2.410		3.6
			20	2.250		3.3	1.615		1.8	1.674		1.8	2.399		3.1
			30	2.264		2.7	1.621		1.5^-	1.668		1.4	2.386		2.6
			60	2.282		1.9	1.628		1.0	1.661		1.0	2.369		1.8
				—	—	+	—	—	+	+	+	+	+	+	+
8.2	0.298	3.354	5	2.141	2.326	8.0	1.571	1.645	4.5^-	1.713	1.645	4.1	2.501	2.326	7.5^+
			10	2.196		5.6	1.594		3.1	1.693		2.9	2.451		5.4
			15	2.221		4.5^+	1.604		2.5^-	1.684		2.4	2.428		4.4
			20	2.235		3.9	1.609		2.2	1.679		2.1	2.415		3.8
			30	2.252		3.2	1.616		1.8	1.673		1.7	2.399		3.1
			60	2.274		2.2	1.625		1.2	1.665		1.2	2.378		2.2
				—	—	+	—	—	+	+	+	+	+	+	+
10.0	0.407	3.570	5	2.114	2.326	9.1	1.559	1.645	5.2	1.724	1.645	4.8	2.533	2.326	8.9
			10	2.176		6.4	1.585		3.6	1.701		3.4	2.473		6.3
			15	2.204		5.2	1.596		3.0	1.691		2.8	2.446		5.2
			20	2.221		4.5^+	1.603		2.6	1.685		2.4	2.430		4.5^-
			30	2.240		3.7	1.611		2.1	1.677		1.9	2.411		3.7
			60	2.265		2.6	1.621		1.5^-	1.668		1.4	2.387		2.6

TABLE 4a. : Errors in t_a.

m	β_1	β_2	n	Lower 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
				—	—	—	—	—	—	+	+	—	+	+	—
2.8	0.056	2.762	5	4.23	3.747	13	2.32	2.132	8	2.03	2.132	5	3.59	3.747	4
			10	3.06	2.821	9	1.94	1.833	6	1.76	1.833	4	2.71	2.821	4
			15	2.80	2.624	7	1.84	1.761	4	1.70	1.761	3	2.52	2.624	4
			20	2.68	2.539	6	1.80	1.729	4	1.68	1.729	3	2.45	2.539	4
			30	2.57	2.462	4	1.75	1.699	3	1.65	1.699	3	2.39	2.462	3
			60	2.47	2.391	3	1.71	1.671	2	1.64	1.671	2	2.34	2.391	2
				—	—	—	—	—	—	+	—	—	+	+	—
2.6	0.100	2.818	5	4.38	As above	17	2.39	As above	12	1.99	As above	7	3.51	As above	6
			10	3.14		11	1.98		8	1.73		6	2.66		6
			15	2.85		9	1.87		6	1.68		5	2.49		5
			20	2.73		8	1.82		5	1.66		4	2.42		5
			30	2.60		6	1.77		4	1.64		3	2.36		4
			60	2.49		4	1.72		3	1.63		2	2.32		3
				—	—	—	—	—	—	+	—	—	+	+	—
2.3	0.207	2.966	5	4.67	As above	25	2.52	As above	18	1.92	As above	10	3.37	As above	10
			10	3.28		16	2.05		12	1.69		8	2.58		9
			15	2.96		13	1.92		9	1.65		6	2.43		7
			20	2.81		11	1.86		8	1.63		6	2.37		7
			30	2.67		8	1.80		6	1.62		5	2.32		6
			60	2.53		6	1.74		4	1.61		4	2.29		4
				—	—	—	—	—	—	+	—	—	+	+	—
1.9	0.492	3.384	5	5.24	As above	40	2.76	As above	29	1.80	As above	16	3.14	As above	16
			10	3.55		26	2.18		19	1.62		12	2.45		13
			15	3.15		20	2.02		15	1.59		10	2.32		12
			20	2.96		17	1.94		12	1.58		9	2.28		10
			30	2.79		13	1.86		9	1.58		7	2.25		9
			60	2.61		9	1.78		7	1.58		5	2.24		6
				—	—	—	—	—	—	+	—	—	+	+	—
1.7	0.748	3.772	5	5.65	As above	51	2.94	As above	38	1.73	As above	19	2.99	As above	20
			10	3.75		33	2.28		24	1.58		14	2.36		16
			15	3.29		25	2.09		19	1.56		11	2.26		14
			20	3.08		21	2.00		16	1.55		10	2.23		12
			30	2.87		17	1.91		12	1.55		9	2.21		10
			60	2.66		11	1.81		8	1.57		6	2.21		8

TABLE 4b. : Errors in t_a.

m	β_1	β_2	n	Lower 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
				—	—	—	—	—	—	+	+	—	+		—
4.9	0.058	2.865	5	4.14	3.747	10	2.31	2.132	8	1.97	2.132	8	3.44	3.747	8
			10	3.03	2.821	7	1.93	1.833	5	1.74	1.833	5	2.66	2.821	6
			15	2.78	2.624	6	1.84	1.761	4	1.69	1.761	4	2.50	2.624	5
			20	2.66	2.539	5	1.79	1.729	4	1.67	1.729	3	2.43	2.539	4
			30	2.56	2.462	4	1.75	1.699	3	1.65	1.699	3	2.38	2.462	3
			60	2.45	2.391	2	1.70	1.671	2	1.64	1.671	2	2.34	2.391	2
				—	—	—	—	—	—	+	+	—	+	+	—
5.5	0.101	2.957	5	4.24	As above	13	2.37	As above	11	1.93	As above	9	3.34	As above	11
			10	3.10		10	1.97		7	1.71		7	2.60		8
			15	2.83		8	1.86		6	1.67		5	2.46		6
			20	2.71		7	1.81		5	1.65		5	2.40		5
			30	2.59		5	1.76		4	1.64		3	2.35		5
			60	2.48		4	1.71		2	1.63		2	2.32		3
				—	—	—	—	—	—	+	+	—	+	+	—
6.9	0.207	3.172	5	4.44	As above	18	2.47	As above	16	1.86	As above	13	3.20	As above	15
			10	3.23		14	2.03		11	1.67		9	2.52		11
			15	2.93		12	1.92		9	1.64		7	2.39		9
			20	2.79		10	1.86		8	1.62		6	2.34		8
			30	2.65		8	1.80		6	1.61		5	2.31		6
			60	2.52		5	1.73		4	1.61		4	2.29		4
				—	—	—	—	—	—	+	+	—	+	+	—
8.2	0.298	3.354	5	4.58	As above	22	2.56	As above	20	1.83	As above	14	3.14	As above	16
			10	3.32		18	2.08		13	1.65		10	2.48		12
			15	3.00		14	1.95		11	1.62		8	2.35		10
			20	2.85		12	1.89		9	1.60		7	2.31		9
			30	2.70		10	1.82		7	1.60		6	2.28		7
			60	2.55		7	1.75		5	1.60		4	2.27		5
				—	—	—	—	—	—	+	+	—	+	+	—
10.0	0.407	3.570	5	4.75	As above	27	2.65	As above	24	1.81	As above	15	3.11	As above	17
			10	3.43		22	2.14		17	1.63		11	2.44		14
			15	3.08		17	2.00		14	1.60		9	2.32		12
			20	2.92		15	1.93		12	1.59		8	2.28		10
			30	2.75		12	1.85		9	1.59		6	2.25		9
			60	2.58		8	1.77		6	1.59		5	2.25		6

TABLE 5a. : Errors in $s^2/\sigma^2 = \chi_a^2/\nu$

m	β_1	β_2	n	Lower 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
						+			+			—			—
2.8	0.056	2.762	5	0.08	0.074	—	0.18	0.178	—	2.33	2.372	2	3.19	3.319	4
			10	0.24	0.232	—	0.38	0.369	2	1.85	1.880	2	2.33	2.407	3
			15	0.35	0.333	5	0.48	0.469	2	1.66	1.692	2	2.02	2.082	3
			20	0.41	0.402	2	0.55	0.532	3	1.56	1.586	2	1.86	1.905	2
			30	0.50	0.492	2	0.62	0.611	1	1.45	1.467	1	1.67	1.710	2
			60	0.63	0.622	1	0.73	0.717	2	1.31	1.321	1	1.45	1.477	2
						+			+			—			—
2.6	0.100	2.818	5	0.08	As above	—	0.18	As above	—	2.34	As above	1	3.24	As above	2
			10	0.24		—	0.38		2	1.86		1	2.36		2
			15	0.34		2	0.48		2	1.67		1	2.05		2
			20	0.41		2	0.54		2	1.57		1	1.88		1
			30	0.50		2	0.62		1	1.45		1	1.69		1
			60	0.63		1	0.72		—	1.31		1	1.46		1
						—			—			+			+
2.3	0.207	2.966	5	0.07	As above	—	0.18	As above	—	2.38	As above	—	3.35	As above	1
			10	0.23		2	0.37		—	1.89		1	2.43		1
			15	0.34		—	0.47		—	1.70		—	2.10		1
			20	0.40		—	0.53		—	1.59		—	1.92		1
			30	0.49		—	0.61		—	1.47		—	1.72		1
			60	0.62		—	0.72		—	1.32		—	1.49		1
						—			—			+			+
1.9	0.492	3.384	5	0.07	As above	—	0.17	As above	—	2.47	As above	4	3.62	As above	9
			10	0.22		—	0.35		5	1.96		4	2.61		8
			15	0.32		4	0.45		4	1.76		4	2.24		8
			20	0.39		3	0.51		4	1.65		4	2.04		7
			30	0.47		4	0.59		3	1.52		4	1.81		6
			60	0.60		4	0.70		2	1.36		3	1.54		4
						—			—			+			+
1.7	0.748	3.772	5	0.06	As above	—	0.16	As above	10	2.54	As above	7	3.85	As above	16
			10	0.21		9	0.34		8	2.02		7	2.76		15
			15	0.30		10	0.43		8	1.81		7	2.35		13
			20	0.37		8	0.49		8	1.69		7	2.13		12
			30	0.46		7	0.57		7	1.55		6	1.89		11
			60	0.58		7	0.68		5	1.38		4	1.59		8

TABLE 5b. : Errors in $s^2/\sigma^2 = \chi_a^2/\nu$

m	β_1	β_2	n	Lower 1 % point			Lower 5 % point			Upper 5 % point			Upper 1 % point		
				Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error	Weibull	Normal	% error
						+			+			—			—
4.9	0.058	2.865	5	0.08	0.074	—	0.18	0.178	—	2.33	2.372	2	3.20	3.319	4
			10	0.24	0.232	—	0.38	0.369	2	1.85	1.880	2	2.34	2.407	3
			15	0.34	0.333	—	0.48	0.469	2	1.67	1.692	1	2.03	2.082	2
			20	0.41	0.402	2	0.54	0.532	2	1.56	1.586	2	1.86	1.905	2
			30	0.50	0.492	2	0.62	0.611	1	1.45	1.467	1	1.67	1.710	2
			60	0.63	0.622	1	0.73	0.717	2	1.31	1.321	1	1.45	1.477	2
						+			+						—
5.5	0.101	2.957	5	0.08	As above	—	0.18	As above	—	2.36	As above	—	3.27	As above	1
			10	0.23		—	0.37		—	1.87		—	2.38		1
			15	0.34		—	0.48		2	1.68		—	2.06		1
			20	0.41		2	0.54		2	1.58		—	1.88		1
			30	0.50		2	0.62		1	1.46		—	1.69		1
			60	0.63		1	0.72		—	1.32		—	1.47		—
						—			—			+			+
6.9	0.207	3.172	5	0.07	As above	—	0.17	As above	—	2.42	As above	2	3.46	As above	4
			10	0.22		—	0.36		—	1.92		2	2.50		4
			15	0.33		—	0.46		—	1.73		2	2.15		3
			20	0.39		3	0.52		2	1.62		2	1.96		3
			30	0.48		2	0.60		2	1.49		2	1.75		2
			60	0.61		2	0.71		—	1.34		1	1.50		2
						—			—			+			+
8.2	0.298	3.354	5	0.07	As above	—	0.17	As above	—	2.49	As above	5	3.65	As above	10
			10	0.22		—	0.35		5	1.97		5	2.61		8
			15	0.31		7	0.45		4	1.77		5	2.24		8
			20	0.38		5	0.51		4	1.65		4	2.04		7
			30	0.47		4	0.59		3	1.52		4	1.81		6
			60	0.60		4	0.70		2	1.36		3	1.54		4
						—			—			+			+
10.0	0.407	3.570	5	0.06	As above	19	0.16	As above	10	2.58	As above	9	3.91	As above	18
			10	0.20		14	0.33		11	2.04		9	2.78		15
			15	0.30		10	0.43		8	1.83		8	2.37		14
			20	0.36		10	0.49		8	1.70		7	2.14		12
			30	0.45		9	0.57		7	1.56		6	1.89		11
			60	0.58		7	0.68		5	1.39		5	1.59		8

(iii) If we are satisfied when the true u-factor differs not more than 5% in value from the « normal theory » value, then this limit of error is scarcely exceeded throughout the whole range of Tables 3 a and b for samples of $n = 15$ or more. For $n = 10$, this limit is only just passed when $\beta_1 = 0.30$.

While these results can be regarded as very satisfactory, the case where σ is known is less important than that in which σ must be estimated by the sample s, and the factor t_a has to be used.

The case of t; Tables 4 a and b.

The situation illustrated in these tables is very different.

(i) The errors are now in the opposite sense, as the distribution of t is negatively skew.

(ii) While there is some correspondence between the upper and lower 5 percent point errors, those for the lower 1 percent points are considerably greater than for the upper 1 percent point. Take, for example, the 98 percent confidence interval for $\alpha = 0.01$ and for samples of 5 from a Weibull population with $m = 2.3$, ($\beta_1 = 0.207$, $\beta_2 = 2.966$). « Normal theory » suggests that the width of interval is

$$2 \times 3.747 \times s/\sqrt{5} = 7.49\ s/\sqrt{5}\,,$$

but the true value is

$$(4.67 + 3.37)\ s/\sqrt{5} = 8.04\ s/\sqrt{5}\,.$$

(iii) At what magnitude the errors involved in using the « normal theory » t-factors becomes serious is, of course, a matter of opinion, but in the majority of cases illustrated they are more, and often very much more, than 5 percent.

The case of $(s/\sigma)^2 = \chi^2/v$; Tables 5 a and b.

(i) From the confidence interval point of view we are concerned with errors in $\sqrt{v}/\chi_\alpha$, expressed as a percentage of the « normal theory » value ; these errors will be roughly one-half of those for χ_α^2/v which have been tabled. The position disclosed is therefore, perhaps, unexpectedly satisfactory; on the upper branch of the Weibull line errors of 10% in χ^2/v (or 5% in $\sqrt{v/\chi}$) have not been reached when β_1 is as large as 0.492. It is certain, however, that the position will deteriorate for populations with betas lying in the area *below* (in the sense of figure 1) the gamma or Type III line. This effect is in fact showing itself when we notice that for a distribution with $m = 10.0$ ($\beta_1 = 0.407$, $\beta_2 = 3.570$) the errors are larger than for one with $m = 1.7$ ($\beta_1 = 0.748$, $\beta_2 = 3.772$).

(ii) The errors for the lower and upper limits are now of opposite sign. As a result for $\beta_1 \leqslant 0.1$ the true confidence limits are a little nearer together than those given by the usual χ^2 factors, and they are further apart for $\beta_1 \geqslant 0.2$.

4. General comment.

The information provided in Tables 3-5 is admittedly limited and to complete the picture similar comparisons are needed for samples from populations whose beta values lie well outside what may be termed the Weibull area. If we consider the distributions given in Table 1 for which the (β_1, β_2) points are plotted in figure 1, the following seems to be the position:

Distributions 1-4. « Normal theroy » should certainly hold good.

Distributions 5-8. The results for the Weibull distributions suggest that no very serious errors would follow in the cases using the « normal theory » factors.

Distributions 9 and 10. Tables 4 a and b suggest the kind of errors involved in using the « normal theory » t values. The errors for u and $(s/\sigma)^2$ will be small, since $\beta_1 \sim 0.20$.

Distribution 11 is a symmetrical, flat topped or platykurtic one, and for this type of population more information seems needed.

Distributions 12 and 13. An idea of what happens for these two rather asymmetrical distributions is provided by the Weibull distribution results for $m = 1.7$, $\beta_1 = 0.748$.

Distributions 14 and 15. More information is certainly needed in the case of these leptokurtic distributions, indeed for populations with β_1, β_2 points in the area below the gamma line. It is likely that for distribution 14, a logarithmic transformation would bring the variable close to normality.

Since this paper was originally drafted, the author has been following up the line of approach suggested in section 2 above, in regard to the distribution of $(s/\sigma)^2$. Although the computations planned have not yet been completed, it seems that it will be possible to draw rough contours in the population (β_1, β_2) plane, appropriate for certain specified sample sizes, within which the percentage error involved in using confidence limits for σ based on the χ-distribution is less than a specified amount, say 5 or 10%.

The main population distributions used will be represented by Pearson curves, but in certain cases a comparison is being made between results for a Pearson Type I [as in equation (6)] and a Weibull distribution [equation (4)] having the same values of (β_1, β_2). The problem in regard to Student's t is less amenable to this procedure as it is clear, for example, that for samples of size $n = 10$, the expansions for the moments of t, as far as they have been taken, do not converge.

REFERENCES

FRIDSHAL, D. & POSTEN, H.O. (1966). *Feasibility study on the determination of regions of robustness in the Pearson system by Monte Carlo techniques.* University of Connecticut, Department of Statistics, Research Report No. 23.

GEARY, R.C. (1947). Testing for normality. *Biometrika,* Vol. 34, pp. 209-42.

HARTER, H.L. & DUBEY, S.D. (1967). Theory and tables for tests of hypotheses concerning the mean and the variance of a Weibull distribution. Report A.R.L. 67-0059. Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio.

JOHNSON, N.L., NIXON, Eric, AMOS, D.E. & PEARSON, E.S. (1963). Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure. *Biometrika,* Vol. 50, pp. 459-98.

KENDALL, Maurice G. & STUART, Alan (1958). *The Advanced Theory of Statistics,* Vol. 1. London : Chas. Griffin and Co. Ltd.

INFLUENCE DE L'ARRET D'UN VEHICULE SUR LE FLOT EN AMONT

par M. ROUBENS
Faculté Polytechnique de Mons

1. Introduction et résumé.

Le comportement individuel des véhicules doit être considéré dans de nombreux problèmes de trafic et plus particulièrement dans les problèmes de simulation de circulation urbaine à l'approche d'un carrefour sans feux. La simulation sur ordinateur nécessite dès lors la génération de véhicules. Deux options sont possibles pour définir le pas horlogique : le pas intervéhiculaire et le pas unitaire classique. Dans le premier cas, la « naissance » d'un véhicule est déterminée par l'intervéhicule, c'est-à-dire par la distance aléatoire (exprimée en unités de temps) séparant deux mobiles. Cet intervéhicule est engendré par tirage dans une loi de probabilité quelconque (exponentielle, gamma, Weibull, ...). Dans le second cas, il est possible, à chaque pas de l'horloge, de créer, avec une probabilité donnée, un nouveau véhicule par tirage dans une loi uniforme. Un schéma de Bernoulli est ainsi constitué par l'ordinateur. Un succès est identifié à la « naissance » d'un véhicule. Dans les deux cas envisagés, une contrainte de distance minimale peut être introduite.

Nous présentons ces deux types de génération, en nous limitant pour le premier cas aux intervéhicules exponentiels (loi d'apparition poissonnienne). Pour le second cas, la distribution de la somme des intervéhicules est calculée et permet de déterminer l'influence de l'arrêt d'un véhicule sur le comportement des suiveurs ainsi que le retard moyen occasionné à ceux-ci.

2. Génération de véhicules par pas aléatoires : intervéhicule exponentiel.

Soient 1, 2, ..., les instants d'apparition des véhicules distribués suivant une loi de Poisson et i_k, les intervéhicules $(k, k+1)$. La fonction de fréquence de i_k, $k = 1, 2, \ldots$, s'écrit

$$f(t) = \begin{cases} \lambda\, e^{-\lambda(t-d_m)} & \text{si} \quad d_m \leqslant t, \\ 0 & \text{si} \quad t < d_m, \end{cases} \tag{1}$$

si λ est le taux d'arrivées des véhicules exprimé en [véhicules/temps] et d_m, la distance intervéhiculaire minimale exprimée en [temps].

La moyenne des intervéhicules vaut

$$E(i_k) = d_m + \frac{1}{\lambda}. \quad (2)$$

Pour engendrer des intervéhicules, il suffit de tirer dans la loi

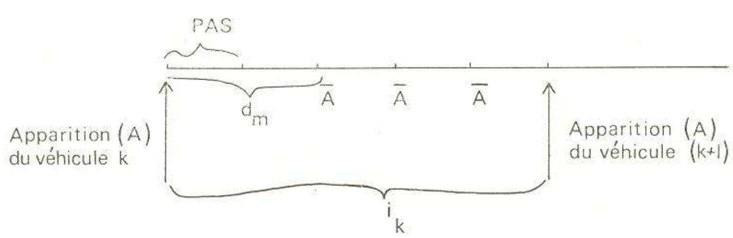
$$F(x) = \int_{d_m}^{x} f(t)\, dt, \qquad x \geqslant d_m.$$

Si y est une variable uniforme sur le segment [0, 1], obtenue à l'aide d'une des règles classiques de génération de nombres au hasard, nous trouvons un intervéhicule i par la formule

$$i = d_m - \frac{1}{\lambda} \log_e y.$$

3. Génération de véhicules par pas unitaires : schéma de Bernoulli.

Considérons le schéma de Bernoulli suivant :



Si p = P (apparition d'un véhicule), $q = 1 - p$ et si x est exprimé en [pas], la distribution des intervéhicules i_k est géométrique et

$$P(i_k = x) = \begin{cases} p\, q^{x-d_m} & \text{si} \quad d_m \leqslant x, \\ 0 & \text{si} \quad x < d_m. \end{cases} \quad (3)$$

$$E(i_k) = d_m + \frac{q}{p} = d_m - 1 + \frac{1}{p}. \quad (4)$$

La formule (4) permet de déterminer la valeur de p correspondant à une distance minimale et un intervéhicule moyen donnés.

La génération sur ordinateur se fait très simplement (voir aussi [1]). A chaque pas, on vérifie si la distance séparant ce pas de la dernière apparition est inférieure, égale ou supérieure à d_m. Lorsque la distance d_m est atteinte ou dépassée, en engendre y, variable uniforme sur [0, 1].

$$\text{Si} \begin{cases} y \leqslant p, & \text{il y a une nouvelle apparition (A),} \\ y > p, & \text{il n'y a pas d'apparition } (\overline{A}). \end{cases}$$

Soit $J_N = \sum_{\nu=1}^{N} j_\nu$, $j_\nu = i_\nu - d_m$. Recherchons $P(J_N = x) = P_N(x)$, les intervéhicules j_ν étant indépendants.

La fonction génératrice de $P_1(x)$, notée $P_1^*(z)$, vaut, par (2),

$$P_1^*(z) = \sum_{x=0}^{\infty} z^x \, P(j_1 = x) = \sum_{x=0}^{\infty} p\,(qz)^x = p\,(1 - qz)^{-1}.$$

Dès lors, avec des notations évidentes,

$$P_N^*(z) = [P_1^*(z)]^N = p^N (1 - qz)^{-N}.$$

Par développement en série de $(1 - qz)^{-N}$, il vient :

$$P_N^*(z) = \sum_{x=0}^{\infty} p^N \binom{N+x-1}{x-1} q^x z^x \text{ et finalement,}$$

$$P_N(x) = \binom{N+x-1}{x-1} p^N q^x. \tag{5}$$

4. Influence de l'arrêt d'un véhicule sur le comportement des suiveurs.

Supposons qu'un véhicule (appelé 1) s'arrête durant un temps A_0. Si la distance minimale (exprimée en [pas]) séparant deux véhicules à l'*arrêt* est d_a et si le retard infligé aux suiveurs pour démarrage vaut R, nous recherchons Q_k, la probabilité pour que k véhicules (notés 2, ..., $k+1$) soient influencés par l'arrêt du véhicule 1.

La distance minimale séparant deux véhicules en *mouvement* vaut d_m ($d_m > d_a$) et la distribution des intervéhicules est géométrique.

A titre d'exemple, considérons $A_0 = 5$ sec, $d_a = 2$ sec, $R = 2$ sec, $d_m = 5$ sec. Si $i_1 = 5$ sec, $i_2 = 6$ sec, $i_3 = 5$ sec, $i_4 = 6$ sec, ..., la figure 1 résume la situation.

Nous constatons, dans ce cas particulier, que les intervéhicules (1, 2), (2, 3), (3, 4) des véhicules 2, 3, 4 en mouvement après l'arrêt sont inférieurs à $d_m = 5$ sec. Une telle situation se présente pour les véhicules influencés par l'arrêt de 1, si $d_m < d_a + R$. Toutefois, la situation réelle s'apparentant à la figure 2, la solution théorique proposée est valable si « influence » signifie arrêt du véhicule et non ralentissement de celui-ci.

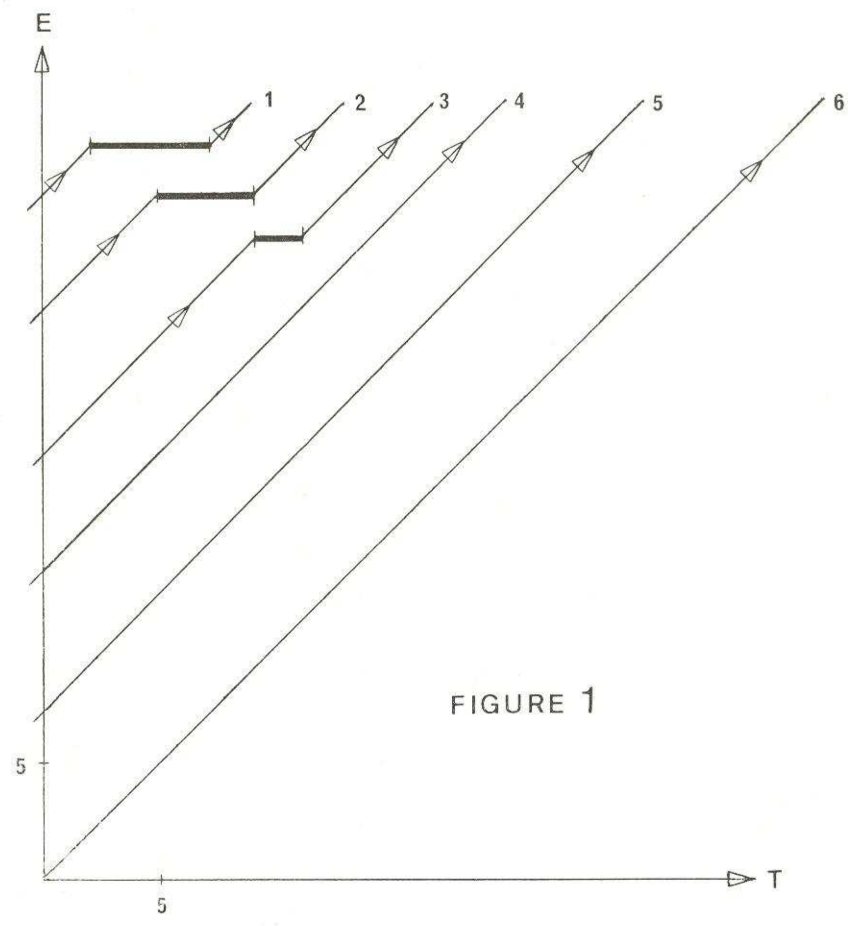


Fig. 1.

L'arrêt du véhicule 1 influence k véhicules si et seulement si :

$$\begin{cases} i_1 \leqslant A_0 + d_a, \\ i_1 + i_2 \leqslant A_0 + 2\,d_a + R, \\ \vdots \\ i_1 + i_2 + \ldots + i_k \leqslant A_0 + k\,d_a + (k-1)\,R, \\ i_1 + i_2 + \ldots + i_{k+1} > A_0 + (k+1)\,d_a + k\,R. \end{cases} \tag{6}$$

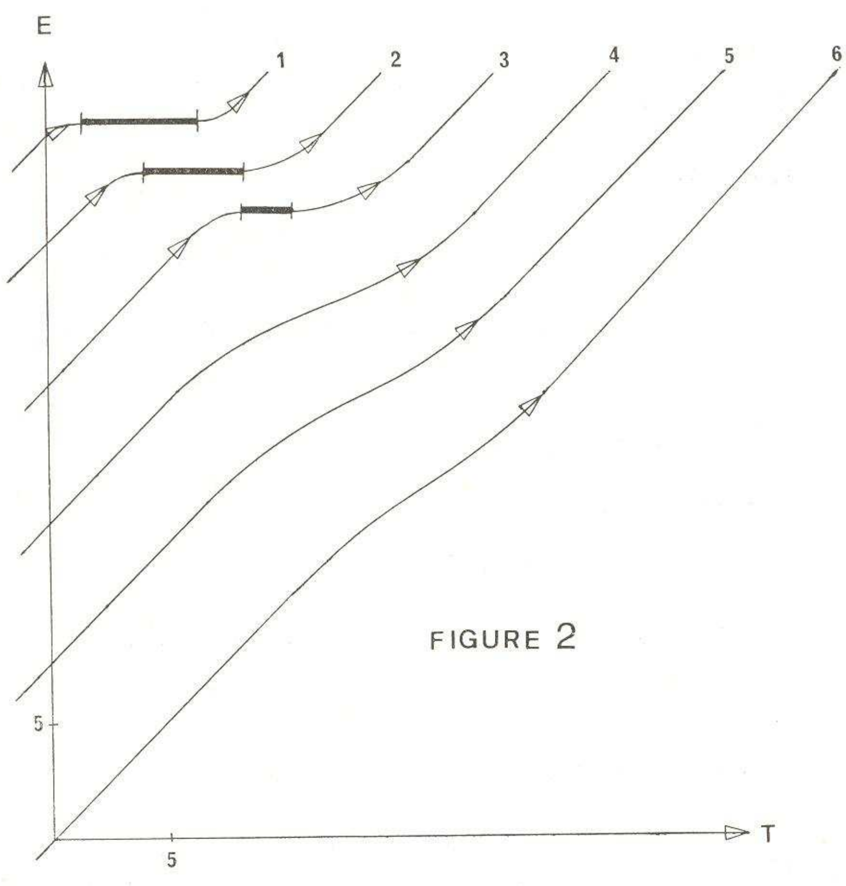


Fig. 2.

$$\text{Si } j_k = i_k - d_m, \qquad J_N = \sum_{\nu=1}^{N} j_\nu, \qquad \theta = d_a - d_m + R,$$

$A_1 = A_0 + d_a - d_m$, $A_k = A_1 + (k-1)\,\theta$, les inégalités (6) s'écrivent :

$$\begin{cases} J_1 \leqslant A_1, \\ J_2 \leqslant A_2, \\ \vdots \\ J_k \leqslant A_k, \\ J_{k+1} > A_{k+1}. \end{cases} \tag{7}$$

Le problème posé ne présente pas d'intérêt si $d_m > d_a + A_0$. En effet, dans ce cas, $A_1 < 0$ et $Q_0 = P(J_1 > A_1) = 1$.

Supposons, dès lors, $d_m \leqslant A_0 + d_a$ et considérons séparément $\theta < 0$ et $\theta > 0$.

a) $\theta < 0$.

Soit r tel que $A_1 > A_2 > \ldots > A_r \geqslant 0$, $A_{r+j} < 0$, $j > 0$.

Dès lors, $Q_{r+j} = P(J_1 \leqslant A_1, \ldots, J_{r+1} \leqslant A_{r+1}, \ldots) = 0$, car l'événement $J_{r+1} \geqslant A_{r+1}$ est impossible.

$$Q_0 = P(J_1 > A_1) = p \sum_{x=A_1+1}^{\infty} q^x = q^{A_1+1},$$

$$\begin{aligned} Q_r &= P(J_1 \leqslant A_1, \ldots, J_r \leqslant A_r, J_{r+1} > A_{r+1}), \\ &= P(J_1 \leqslant A_1, \ldots, J_r \leqslant A_r), \text{ car } A_{r+1} < 0, \\ &= P(J_r \leqslant A_r) = 1 - \sum_{i=0}^{r-1} Q_i = \sum_{x=0}^{A_r} \binom{x+r-1}{r-1} p^r q^x. \end{aligned}$$

Il reste donc à calculer Q_j, si $0 < j < r$.

$$Q_j = P(J_1 \leqslant A_1, \ldots, J_j \leqslant A_j, J_{j+1} > A_{j+1}).$$

Comme $J_j \leqslant A_j \rightarrow \begin{cases} J_1 \leqslant A_1, & (A_1 > A_2 > \ldots > A_j), \\ \vdots \\ J_{j-1} \leqslant A_{j-1}. \end{cases}$ nous obtenons

$$Q_j = P(J_j \leqslant A_j\,,\ J_{j+1} > A_{j+1})\,,$$

$$= P(J_j \leqslant A_{j+1}\,,\ J_{j+1} > A_{j+1}) + P(A_{j+1} < J_j \leqslant A_{j+1})\,,$$

$$= \sum_{x=0}^{A_{j+1}} \binom{x+j-1}{j-1} p^j q^x\, P(j_{j+1} > A_{j+1} - x) + \sum_{x=A_{j+1}+1}^{A_j} \binom{x+j-1}{j-1} p^j q^x\,,$$

en vertu de (5)

$$= \sum_{x=0}^{A_{j+1}} \binom{x+j-1}{j-1} p^j q^x q^{A_{j+1}-x+1} + \ldots .$$

Une démonstration par récurrence conduit à

$$\sum_{\nu=0}^{\alpha} \binom{\nu+\beta}{\beta} = \binom{\alpha+\beta+1}{\beta+1}.$$

Cette dernière formule permet d'écrire finalement :

$$\begin{cases} Q_0 = q^{A_1+1} \\ Q_j = p^j \left\{ \binom{A_{j+1}+j}{j} q^{A_{j+1}+1} + \sum_{x=A_{j+1}+1}^{A_j} \binom{x+j-1}{j-1} q^x \right\}, \quad 0 < j < r; \\ Q_r = 1 - \sum_{i=0}^{r-1} Q_i = \sum_{x=0}^{A_r} \binom{x+r-1}{r-1} q^x p^r\,, \\ Q_{r+j} = 0\,,\ j > 0\,. \end{cases} \tag{8}$$

b) $\theta \geqslant 0$.

Dans ce cas, $A_1 \leqslant A_2$, ..., et $Q_j \neq 0$, pour tout $j \geqslant 0$.

$$Q_0 = P(J_1 > A_1) = Q^{A_1+1}\,.$$

$$Q_j = P(J_1 \leqslant A_1\,, \ldots, J_j \leqslant A_j\,,\ J_{j+1} > A_{j+1}),\ j > 0\,,$$

$$= \sum_{x_1=0}^{A_1} \cdots \sum_{x_j=0}^{A_j - x_1 - \ldots - x_{j-1}} \ \sum_{x_{j+1}=A_{j+1}-x_1-\ldots-x_j+1}^{\infty} p^{j+1} q^{x_1+\ldots+x_{j+1}}$$

$$= \sum_{x_1=0}^{A_1} \cdots \sum_{x_j=0}^{A_j-\dots-x_{j-1}} p^j q^{x_1+\dots+x_j} q^{A_{j+1}-x_1-\dots-x_j+1}$$

$$= p^j Q^{A_{j+1}+1} \{ \sum_{x_1=0}^{A_1} \cdots \sum_{x_j=0}^{A_j-x_1-\dots-x_{j-1}} 1 \} .$$

Pour faciliter le calcul de la série multiple, il suffit de poser :

$$y_k = A_k - x_1 - \dots - x_k \quad \text{et, dès lors,}$$

$$\begin{cases} Q_0 = q^{A_1+1} , \\ Q_j = p^j q^{A_{j+1}+1} \{ \sum_{y_1=0}^{A_1} \sum_{y_2=0}^{y_1+\theta} \cdots \sum_{y_j=0}^{y_{j-1}+\theta} 1 \} \; j > 0 . \end{cases} \tag{9}$$

Nous obtenons ainsi,

$$\begin{cases} Q_0 = q^{A_1+1} , \\ Q_1 = (A_1 + 1) \; p \; q^{A_2+1} , \\ Q_2 = \dfrac{(A_1 + 1) (A_3 + 2)}{2} \; p^2 \; q^{A_3+1} , \\ \vdots \end{cases}$$

Si les intervéhicules sont exponentiels, la probabilité Q_k s'écrit (voir [2]),

$$Q_k = \int_0^{A_1} \lambda e^{-\lambda j_1} \, d j_1 \dots \int_0^{A_k-j_1-\dots-j_{k-1}} \lambda e^{-\lambda j_k} \, d j_k \int_{A_{k+1}-j_1-\dots j_k}^{\infty} \lambda e^{-\lambda j_{k+1}} \, d \, j_{k+1} .$$

Par le changement de variables $y_k = A_k - J_k$, $k > 0$, nous obtenons :

$$Q_k = \lambda^k \; e^{-\lambda A_{k+1}} \int_0^{A_1} dy_1 \int_0^{y_1+\theta} dy_2 \dots \int_0^{y_{k-1}+\theta} dy_k .$$

On montre par récurrence que :

$$\int_0^{y_1+\theta} dy_2 \dots \int_0^{y_{r-1}+\theta} dy_r = \frac{1}{(r-1)\,!} \; (y_1 + \theta) \; (y_1 + r\theta)^{r-2} .$$

$$\text{Dès lors,} \quad \begin{cases} Q_0 = e^{-\lambda A_1} \\[2ex] Q_k = \dfrac{\lambda^k}{k\,!}\, e^{-\lambda A_{k+1}}\, A_1\, A_{k+1}^{k-1}, \quad k > 0. \end{cases} \tag{10}$$

Les formules (2) et (10) permettent d'obtenir des approximations continues (intervéhicules exponentiels) des formules (4) et (9) du problème discret (intervéhicules géométriques), si les conditions suivantes sont respectées :

— d_m (discret) $\geqslant 1$, (11)

— d_m (discret) $= d_m$ (continu) $+ 1$, (12)

— $p = \lambda$, (13)

— p proche de zéro ($p \leqslant .1$) (14)

En effet, d_m (discret) ne peut être nul afin d'éviter une accumulation de véhicules en un point. De plus, la relation (12) permet — par comparaison de (2) et (4) — d'écrire (13), relation intuitivement évidente. Enfin, l'approximation de la distribution géométrique par la distribution exponentielle n'est valable que si p est proche de zéro (dans ce cas $e^{-p} \cong 1 - p$).

Considérons le cas particulier où

$$A_0 = 2, \quad d_a = 1, \quad R = 1,$$
$$d_m = 0, \quad \lambda = .1,$$
$$d_m = 1, \quad p = .1.$$

Nous obtenons les résultats suivants :

Discret	Continu
$A_1 = 2,\ A_2 = 3,\ A_3 = 4, \ldots$	$A_1 = 3,\ A_2 = 5,\ A_3 = 7, \ldots$
$Q_0 = .729$	$Q_0 = .741$
$Q_1 = .197$	$Q_1 = .182$
$Q_2 = .053$	$Q_2 = .057$

5. Arrêt moyen infligé par l'arrêt du véhicule 1.

Si le véhicule $(j + 1)$ est influencé par l'arrêt du véhicule 1, $J_j \leqslant A_j$ et la durée d'arrêt, notée a_{j+1}, vaut

$$a_{j+1} = A_j + R - J_j \, .$$

Appelons $\bar{a}_{j+1}$, la durée moyenne d'arrêt du véhicule $(j + 1)$.

$$\bar{a}_{j+1} = \sum_{k=0}^{A_j} (A_j + R - k) \binom{k+j-1}{j-1} p^j q^k \qquad \text{[voir (5)]}$$

Si k véhicules sont influencés par l'arrêt A_0, nous notons

$$\overline{(A/k)} = \frac{1}{k} \sum_{j=1}^{k} \bar{a}_{j+1}$$

et, dès lors, le retard moyen infligé par l'arrêt A_0 du véhicule 1 vaut :

$$\text{si } \theta < 0, \; \overline{A} = \sum_{i=1}^{k} \overline{(A/i)} \, Q_i \, , \qquad (A_k \geqslant 0, \; A_{k+1} < 0) \, ,$$

$$\text{si } \theta \geqslant 0, \; \overline{A} = \sum_{i=1} \overline{(A/i)} \, Q_i \, .$$

REFERENCES

[1] R.M. LEWIS : A proposed headway distribution for traffic simulation studies. Traffic Engineering, Vol. 33, n° 5 (1963), pp. 16-19.

[2] Ph. PASSAU : Files d'attente et processus markoviens à une intersection de trafic. In « Queuing Theory - Recent Developments and Applications », The English Universities Press, London, 1967, pp. 185-192.

PUBLICATIONS REÇUES — ONTVANGEN PUBLICATIES

1) Revue française d'Informatique et de Recherche Opérationnelle, AFCET, Paris, N°s 14 et 15, 1968.

2) Annales de Sciences Economiques Appliquées, N° 2, mai 1969.

3) Cuadernos de Estadistica Aplicada e Investigación Operativa, Barcelone, Vol. VI, Fasc. I, 1969.

4) Quality Progress, American Society for Quality Control, Vol. I, N° 8.

5) Revue IBN (Institut Belge de Normalisation), N°s 4 et 5, 1969.
BIN Revue (Belgisch Instituut voor Normalisatie), Nrs 4 en 5, 1969.

6) Revue Suisse pour l'Organisation Industrielle (Institut d'Organisation Industrielle de l'Ecole Polytechnique Fédérale de Zurich), N°s 5 et 6, 1969.

7) Project Cost Control using Networks (The Operational Research Society & the Institute of Cost & Works Accountants), London.

8) AGIFORS PROCEEDINGS VIII (Proceedings of the Eighth AGIFORS Symposium) October 7-9, 1968, Princeton, U.S.A.

9) Revue de Statistique Appliquée, Vol. 17, N° 2, 1969.

10) Operational Research Quaterly, London, Vol. 20, Nr 2, 1969.

11) Revue de l'Institut international de Statistique, La Haye, Vol. 36, N° 2, 1968.